Pure and Applied Functional Analysis
Volume 5, Number 2, 2020, 259–285

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

OPTIMAL OUTPUT FEEDBACK CONTROL LAW FOR PARTIALLY OBSERVED STOCHASTIC EVOLUTION EQUATIONS ON UMD-BANACH SPACES

N. U. AHMED

Abstract. In this paper we consider a general class of semilinear partially observed stochastic evolution equations on UMD-Banach spaces and their optimal feedback controls. The system is governed by a pair of coupled stochastic evolution equations, one of which represents the main system to be controlled, and the other represents the observer designed to monitor the main system and provide information for control. The coupled system is defined on the Cartesian product of two UMD spaces. We prove existence, uniqueness and regularity properties of (mild) solutions of the coupled system on UMD spaces. Then, using the strong operator topology on the space of bounded linear operators, we introduce the class of feedback control laws given by operator valued functions endowed with the Tychonoff product topology. We prove the existence of optimal feedback control laws. Then we consider the attainable set of induced measure valued functions and prove existence of optimal controls for several interesting measure related control problems. Further, we present also the necessary conditions of optimality and a convergence theorem for an algorithm based on the necessary conditions.

1. Introduction

In recent years we have studied control problems for uncertain dynamic systems on Banach spaces [7, 8] where the uncertainty is characterized merely by the geometry and size of its range not by any particular probabilistic structure. This is particularly important if the dynamic model is not fully known or the system has many unknown parameters. In this paper we consider stochastic systems driven by well defined random processes as opposed uncertain systems. We consider a general class of stochastic evolution equations on UMD Banach spaces [10, 19]. The system consists of a pair of coupled semilinear stochastic differential equations one of which represents the main dynamic system to be controlled which is defined on a Banach space X, and the other represents the dynamics of the measurement system (observer) defined on a Banach space Y (the output space). The later system monitors the state of the main system and provides partial information about the state to the controller through its own state which is fully accessible. The combined system is called partially observed stochastic control system as opposed to fully observed

2010 *Mathematics Subject Classification.* 34G20, 34H05, 49J27,49J55, 49K27,93C20.
Key words and phrases. Stochastic evolution equations, UMD spaces, Existence of solutions, Optimal operator valued functions as controls, Necessary conditions of optimality.

system. The problem is to find a feedback control law that optimizes certain performance objectives which is described by the expected value of certain nonlinear functional of the state and output trajectories. Fully observed control problems are well studied in the literature as seen in [2, 3] and the references therein. There, using the Bellman's principle of optimality, one obtains an HJB equation which is a nonlinear partial differential equation on infinite dimensional Hilbert or Banach spaces depending on the original state space. The optimal control law is then given in terms of the solution of the HJB (Hamilton-Jacobi-Bellman) equation; see [3] for details. Solving HJB equations on infinite dimensional spaces and then constructing the feedback control law using the solution is a formidable task. In the literature on partially observed control problems, one uses a control law which is a function of the estimated state which is adapted to the available information or observed process. This leads to a filtering problem requiring the solution of Zakai equation for conditional un-normalized measure valued functions [2]. Using this measure valued function one obtains the so called filtered (or estimated) version of the state. The optimal control is given by some function of this estimated state. This certainly is an indirect approach requiring solution of partial differential equations on R^n in case of finite dimensional control problems [2, Chapter 15] while, in case of infinite dimensional control problems, one is required to solve partial differential equations defined on infinite dimensional Hilbert or Banach space [3]. Clearly this is also a formidable problem.

Here in this paper we formulate the original infinite dimensional control problem in a natural and direct way where one is required to find the optimal control law as a map from the output space of the monitor to the input space of the system to be controlled. This gives rise to interesting topologies and topological questions such as continuity and compactness.

The rest of the paper is organised as follows. In section 2, we present the system model and formulate the control problem. Basic assumptions are presented in section 3. In section 4, we present existence, uniqueness and regularity properties of solutions of the system equations. In section 5, we consider control problems. Here we introduce a general class of admissible feedback operators and a suitable topology on the space of such operator valued functions. Then we prove the continuity of solutions with respect to control operators in the given topology (Theorem 5.2). Using the topological properties of the admissible class, existence of optimal feedback control laws (operator valued functions) is proved (Theorem 5.3). In section 6, we introduce the attainable set of measures induced by the family of solutions corresponding to the admissible set of Control operators. We prove its weak compactness in the space of probability measures and present several interesting results on optimal controls involving the measures. In section 7, we present the necessary conditions of optimality (Theorem 7.1) whereby one can construct the optimal policy. The last section (Section 8), is devoted to the question of convergence of an algorithm based on the necessary conditions of optimality giving Theorem 8.1.

2. Basic formulation of the system model

Let X, Y be a pair of real Banach spaces and consider the system governed by a pair of interconnected stochastic evolution equations on the Cartesian product

$X \times Y \equiv Z$ as follows:

$$dx = Axdt + F(t,x)dt + B(t)ydt + G(t,x)dW, x(0) = x_0, t \in I, \tag{2.1}$$

$$dy = A_0ydt + F_0(t,x)dt + G_0(t,y)dW_0, y(0) = y_0, t \in I \equiv [0,T]. \tag{2.2}$$

The space X denotes the state space of the system to be controlled given by equation (2.1). The process x is not accessible and so not observable. The space Y denotes the state space of the system described by equation (2.2) called the observer, whose state is accessible and so observable. The observable process y is monitored in the presence of noise and used to control the stochastic system defined by equation (2.1). Both W and W_0 are H and H_0-cylindrical Brownian motions defined on a complete probability space with a filtration $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, P)$. Both the systems are noisy since their initial states $\{x_0, y_0\}$ are random and they are driven by Brownian motions which we assume to be mutually independent. Let $(\mathcal{L}(Y,X), \tau_{so})$ denote the space of bounded linear operators $\mathcal{L}(Y,X)$ furnished with the strong operator topology τ_{so} and let $\mathcal{B}_{ad}$ denote the set of admissible operator valued functions defined on the interval $I \equiv [0,T]$ and taking values in $\Lambda \subset (\mathcal{L}(Y,X), \tau_{so})$, a bounded set with bound $b < \infty$. The problem is to find an element $B \in \mathcal{B}_{ad}$ that minimizes the cost functional

$$J(B) \equiv \mathbf{E}\bigg\{\int_I \ell(t,x,y)dt + \Phi(x(T), y(T))\bigg\}. \tag{2.3}$$

More characterization of the set $\mathcal{B}_{ad}$ will follow shortly. This is the natural setting of many real world problems where there is an underlying large scale distributed system whose state is either not physically accessible or its dimension is so large that it is prohibitively costly to monitor the full state and use the data to control the system. However, the system is partially observable in the sense that the state of the main system affects the state of another dynamic system (monitor or observer) whose state is completely accessible. On the basis of this available and possibly imperfect information, one must control the main system in order to realize certain objectives, for example, minimize the cost functional (2.3).

3. Basic assumptions

Throughout the paper, we assume that the Banach spaces X and Y are UMD Banach spaces. This assumption is essential for stochastic integration of X and /or Y valued random processes with respect to Brownian motions. Let $\{H, H_0\}$ be a pair of real separable Hilbert spaces and $\{W(t), W_0(t), t \in I\}$ a pair of H-cylindrical and H_0-cylindrical Brownian motions respectively and that they are stochastically independent. Let $\mathcal{L}(H,X)$ and $\mathcal{L}(H_0,Y)$ denote the spaces of bounded linear operators from the Hilbert H to the Banach space X, and from the Hilbert space H_0 to the Banach space Y respectively. Let $\gamma(H,X) \subset \mathcal{L}(H,X)$ and $\gamma(H_0,Y) \subset \mathcal{L}(H_0,Y)$ denote the spaces of γ-Radonifying operators. Here we present the basic assumptions used in the paper.

(A1): The operator A is the infinitesimal generator of a C_0-semigroup $S(t), t \geq 0$, on a UMD-type2 Banach space X,

(A2): $F : I\times\Omega\times X \longrightarrow X$ is a Borel measurable map, and there exist constants $C_1, C_2 > 0$ such that

(i): $\| F(t,x) \|_X \leq C_1(1+ \| x \|_X),\ \forall\, x \in X,$

(ii): $\| F(t,x_1) - F(t,x_2) \|_X \leq C_2(\| x_1 - x_2 \|_X)\ \forall\, x_1, x_2 \in X.$

(A3): $G : I\times\Omega\times X \longrightarrow \gamma(H,X) \subset \mathcal{L}(H,X)$ is a Borel measurable map, and there exist constants $C_3, C_4 > 0$ such that

(i): $\| G(t,x) \|^2_{\gamma(H,X)} \leq C_3^2(1+ \| x \|^2_X),\ \forall\, x \in X,$

(ii): $\| G(t,x_1) - G(t,x_2) \|^2_{\gamma(H,X)} \leq C_4^2 \| x_1 - x_2 \|^2_X\ \ \forall\, x_1, x_2 \in X,$

(A4): The operator A_0 is the infinitesimal generator of a C_0-semigroup $S_0(t), t \geq 0$, on a UMD-type2 Banach space Y,

(A5): $F_0 : I\times\Omega\times X \longrightarrow Y$ is a Borel measurable map and there exist constants $C_5, C_6 > 0$ such that

(i): $\| F_0(t,x) \|_Y \leq C_5(1+ \| x \|_X),\ \forall\, x \in X,$

(ii): $\| F_0(t,x_1) - F_0(t,x_2) \|_Y \leq C_6 \| x_1 - x_2 \|_X\ \ \forall\, x_1, x_2 \in X,$

(A6): $G_0 : I\times\Omega\times Y \longrightarrow \gamma(H_0,Y) \subset \mathcal{L}(H_0,Y)$ is a Borel measurable map and there exist constants $C_7, C_8 > 0$ such that

(i): $\| G_0(t,y) \|^2_{\gamma(H_0,Y)} \leq C_7^2(1+ \| y \|^2_Y),\ \forall\, y \in Y,$

(ii): $\| G_0(t,y_1) - G_0(t,y_2) \|^2_{\gamma(H_0,Y)^2} \leq C_8^2 \| y_1 - y_2 \|^2_Y\ \ \forall\, y_1, y_2 \in Y.$

4. Existence and uniqueness of solutions

In this section we consider briefly the question of existence and uniqueness of mild solutions of the stochastic evolution equations (2.1)-(2.2). By definition, the mild solutions of these equations are given by the solutions (if any) of the following integral equations defined on the Cartesian product $\mathcal{Z} \equiv X \times Y$ of the Banach spaces X and Y as follows:

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(s,x(s))ds + \int_0^t S(t-s)B(s)y(s)ds + \int_0^t S(t-s)G(s,x(s))dW(s), t \in I, \tag{4.1}$$

$$y(t) = S_0(t)y_0 + \int_0^t S_0(t-s)F_0(s,x(s))ds + \int_0^t S_0(t-s)G_0(s,y(s))dW_0, t \in I. \tag{4.2}$$

We introduce the operators Γ_1, Γ_2 defined on the Banach space $\mathcal{Z}$ and given by the following expressions:

$$\Gamma_1(x,y)(t) \equiv S(t)x_0 + \int_0^t S(t-s)F(s,x(s))ds \tag{4.3}$$

$$+\int_0^t S(t-s)B(s)y(s)ds + \int_0^t S(t-s)G(s,x(s))dW(s), t \in I,$$

$$\Gamma_2(x,y)(t) \equiv S_0(t)y_0 + \int_0^t S_0(t-s)F_0(s,x(s))ds \tag{4.4}$$
$$+\int_0^t S_0(t-s)G_0(s,y(s))dW_0, t \in I.$$

Thus the question of existence of solutions of the integral equations (4.1)-(4.2) is equivalent to the question of existence of a fixed point of the operator Γ given by $z = \Gamma(z)$ where

$$z = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \Gamma_1(x,y) \\ \Gamma_2(x,y) \end{bmatrix} \equiv \Gamma(z).$$

Let $B^a_\infty(I, L_2(\Omega, X))$ and $B^a_\infty(I, L_2(\Omega, Y))$ denote the Banach spaces of $\mathcal{F}_t$-adapted processes defined on the interval I, taking values in the Banach spaces X and Y respectively and having finite second moments. An element $x \in B^a_\infty(I, L_2(\Omega, E))$ has the norm $\parallel x \parallel_\infty$ given by

$$\parallel x \parallel^2_\infty \equiv \sup\{\mathbf{E} \parallel x(t) \parallel^2_E, t \in I\},$$

for $E = X,$ or Y. For convenience of notation, we set

$$B^a_\infty(I, L_2(\Omega, X)) \times B^a_\infty(I, L_2(\Omega, Y)) \equiv B^a_\infty(I, L_2(\Omega, \mathcal{Z})).$$

We present the following result which we need later in the paper.

Theorem 4.1. *Consider the stochastic systems (2.1)-(2.2) and suppose the basic assumptions* **(A1)-(A6)** *hold. Then, for every pair of $\mathcal{F}_0$-measurable initial states $x_0 \in L_2(\Omega, X)$ and $y_0 \in L_2(\Omega, Y)$, and feedback operator $B \in \mathcal{B}_{ad}$, the system (2.1)-(2.2) has a unique $\mathcal{F}_t$-adapted mild solution with $x \in B^a_\infty(I, L_2(\Omega, X))$ and $y \in B^a_\infty(I, L_2(\Omega, Y))$.*

Proof. The proof is standard and it follows from Banach fixed point theorem applied to the operator Γ. We present only some major points. First consider the stochastic integrals in (4.4) and (4.5). We recall the fact that the γ-Radonifying operators form two sided ideal in the space of bounded linear operators. Thus, considering the stochastic integral in (4.4) and defining

$$V_1(t) \equiv \int_0^t S(t-s)G(s,x(s))dW, t \in I, \tag{4.5}$$

it is easy to verify that

$$\mathbf{E} \parallel V_1(t) \parallel^2_X = \int_0^t \mathbf{E} \parallel S(t-s)G(s,x(s)) \parallel^2_{\gamma(H,X)} ds \tag{4.6}$$
$$\leq M^2 C_3^2 \int_0^t (1 + \mathbf{E} \parallel x(s) \parallel^2) ds.$$

Similarly, considering equation (4.5) and denoting the stochastic integral term by V_2 we obtain

$$\mathbf{E} \parallel V_2(t) \parallel^2_Y = \int_0^t \mathbf{E} \parallel S_0(t-s)G_0(s,y(s)) \parallel^2_{\gamma(H_0,Y)} ds \tag{4.7}$$

$$\leq M_0^2 C_7^2 \int_0^t (1 + \mathbf{E} \parallel y(s) \parallel_Y^2) ds.$$

Now using the expressions (4.4) and (4.5) and standard triangle and Hölder inequalities one can verify that there exist positive constants k_1, k_2 depending on constants $\{(C_1) - (C_6), M_0, M\}$ such that

$$\mathbf{E} \parallel \Gamma_1(x,y)(t) \parallel_X^2 \leq k_1 \Big\{ \mathbf{E} \parallel x_0 \parallel_X^2 + (1+t) \int_0^t \mathbf{E} \parallel x(s) \parallel_X^2 ds + t \int_0^t \mathbf{E} \parallel y(s) \parallel_Y^2 ds \Big\} \tag{4.8}$$

and

$$\mathbf{E} \parallel \Gamma_2(x,y)(t) \parallel_Y^2 \leq k_2 \Big\{ \mathbf{E} \parallel y_0 \parallel_Y^2 + (1+t) \int_0^t \mathbf{E} \parallel x(s) \parallel_X^2 ds + \int_0^t \mathbf{E} \parallel y(s) \parallel_Y^2 ds \Big\}. \tag{4.9}$$

From the above estimates we obtain the following inequality

$$\sup\{\mathbf{E} \parallel \Gamma(z)(t) \parallel_{\mathcal{Z}}^2, t \in I\} \leq \alpha \mathbf{E} \parallel z_0 \parallel_{\mathcal{Z}}^2 + \beta \sup\{\mathbf{E} \parallel z(t) \parallel_{\mathcal{Z}}^2, t \in I\} \tag{4.10}$$

where the constants α, β are dependent on the basic parameters $\{(C1) - (C6), M, M_0, T\}$. Clearly, it follows from the above inequality that the operator

$$\Gamma : B_\infty^a(I, L_2(\Omega, \mathcal{Z})) \longrightarrow B_\infty^a(I, L_2(\Omega, \mathcal{Z})).$$

Using similar computations one can verify that for all $t \in I$,

$$\begin{aligned} &\mathbf{E} \parallel (\Gamma(z_1) - \Gamma(z_2))(t) \parallel_{\mathcal{Z}}^2 \\ &\leq \Big\{ (4M^2(tC_2^2 + C_4^2) + 2tM_0^2 C_6^2) \int_0^t \mathbf{E} \parallel x_1(s) - x_2(s) \parallel_X^2 ds \\ &\quad + (4tM^2 b^2 + 2M_0^2 C_8^2) \int_0^t \mathbf{E} \parallel y_1(s) - y_2(s) \parallel_Y^2 ds \Big\}. \end{aligned} \tag{4.11}$$

It follows from this that there exists a constant $\tilde{C}$ dependent on the parameters as displayed $\hat{C} \equiv \tilde{C}(M, M_0, b, C_2, C_4, C_6, C_8, T)$, such that

$$\mathbf{E} \parallel (\Gamma(z_1) - \Gamma(z_2))(t) \parallel_{\mathcal{Z}}^2 \leq \tilde{C}^2 \int_0^t \mathbf{E} \parallel z_1(s) - z_2(s) \parallel_{\mathcal{Z}}^2 ds, t \in I. \tag{4.12}$$

Hence, for any $t_1 \in I$, it follows from the above inequality that

$$\sup_{0 \leq t \leq t_1} \mathbf{E} \parallel (\Gamma(z_1) - \Gamma(z_2))(t) \parallel_{\mathcal{Z}}^2 \leq (\tilde{C}^2 t_1) \sup_{0 \leq t \leq t_1} \mathbf{E} \parallel z_1(t) - z_2(t) \parallel_{\mathcal{Z}}^2 . \tag{4.13}$$

From here, one can follow either of two ways to prove the existence and uniqueness of solution. Using the inequality (4.12) and considering the iterates of the operator Γ, in particular the n-th iterate Γ^n, one can easily verify that

$$\begin{aligned} &\sup_{t \in I} \{\mathbf{E} \parallel (\Gamma^n(z_1) - \Gamma^n(z_2))(t) \parallel_{\mathcal{Z}}^2 \\ &\leq ((\tilde{C}^2 T)^n / n!) \sup_{t \in I} \mathbf{E} \parallel z_1(t) - z_2(t) \parallel_Z^2 . \end{aligned} \tag{4.14}$$

Hence we conclude that

$$\| \Gamma^n(z_1) - \Gamma^n(z_2) \|_{B^a_\infty(I,L_2(\Omega,\mathcal{Z}))} \leq \alpha_n \| z_1 - z_2 \|_{B^a_\infty(I,L_2(\Omega,\mathcal{Z}))}, \tag{4.15}$$

where $\alpha_n = ((\tilde{C}^2 T)^n/n!)^{1/2}$. Clearly, for n sufficiently large, $\alpha_n < 1$ and hence the operator Γ^n is a contraction and consequently it follows from Banach fixed point theorem that Γ^n has a unique fixed point say $z^o \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$. Hence one can easily verify that z^o is also a unique fixed point of the operator Γ. This is one approach. The second method is based on the inequality (4.13). Here one chooses $t_1 \in I$, sufficiently small, so that $\tilde{C}^2 t_1 = \gamma < 1$. This implies that the operator Γ restricted to the Banach space $B^a_\infty(I_1, L_2(\Omega, \mathcal{Z}))$, where $I_1 \equiv (0, t_1]$, is a contraction and hence has a unique fixed point $\zeta \in B^a_\infty(I_1, L_2(\Omega, \mathcal{Z}))$. Considering the interval $I_2 \equiv (t_1, t_2]$ and starting with the state $\zeta(t_1)$, again one chooses a $t_2 \in (t_1, T]$ so that $\tilde{C}^2(t_2 - t_1) = \gamma$, which implies that the operator Γ, restricted to the Banach space $B^a_\infty(I_2, L_2(\Omega, \mathcal{Z}))$, is a contraction and hence has a unique fixed point. This process can be continued till the interval $I = [0, T]$ is covered. By concatenation of the pieces one obtains a unique solution. This completes the outline of our proof. □

Remark 4.2. Under an additional assumption, following the well known Da-Prato Zabczyk factorization technique [11, Theorem 1.1, p144], one can easily verify that the solution process $z = (x, y)$ has continuous modifications. In other words, $z \in L^a_2(\Omega, C(I, \mathcal{Z}))$.

5. Optimal output feedback control laws

Admissible Set of Feedback Control Operators: To prove existence of optimal output feedback control laws (operators), we use continuous dependence of solutions with respect to such operators. For this we must define appropriate topology on the space of operator valued functions. Consider the space of bounded linear operators $\mathcal{L}(Y, X)$ endowed with the strong operator topology τ_{so} and denote this by $\mathcal{L}_{so}(Y, X)$. It is well known that $\mathcal{L}_{so}(Y, X)$ is a locally convex sequentially complete Hausdorff topological vector space. We choose a τ_{so} compact set $\Lambda \subset \mathcal{L}_{so}(Y, X)$ and let $\mathcal{B}_{ad} \equiv B_0(I, \Lambda)$ denote the class of strongly measurable operator valued functions defined on I and taking values from the set Λ. Here we can use the well known Tychonoff product topology [13] denoted by $\tau_{\tau\pi}$. Endowed with the Tychonoff product topology, $\mathcal{B}_{ad}$ is also a compact Hausdorff space. We choose this as the set of admissible feedback operators.

Remark 5.1. Since the set Λ, defining the admissible set of feedback operators $\mathcal{B}_{ad} \equiv B_0(I, \Lambda)$, is compact in the strong operator topology τ_{so}, it follows from uniform boundedness principle (Banach-Steinhaus theorem) that it is bounded in the uniform operator norm topology. Hence the solution set

$$\mathcal{S} \equiv \{z(B) : z(B) \in B^a_\infty(I, L_2(\Omega, \mathcal{Z})), B \in \mathcal{B}_{ad}\}$$

is a bounded subset of $B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$.

We prove the following continuity result.

Theorem 5.2. *Consider the system (2.1)-(2.2) with the admissible control operators $\mathcal{B}_{ad}$, and suppose the assumptions of Theorem 4.1 hold. Then, the map*

$B \longrightarrow (x,y) \equiv z$ is continuous with respect to the Tychonoff product topology $\tau_{\tau\pi}$ on $\mathcal{B}_{ad}$ and the natural norm topology on $B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$.

Proof. Let $\{B_n, B_o\} \in \mathcal{B}_{ad}$ and suppose $B_n \xrightarrow{\tau_{\tau\pi}} B_o$ (in the Tychnoff product topology). Let $\{z_n\} \equiv \{(x_n, y_n)\} \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ denote the mild solutions of the system (2.1)-(2.2) corresponding to the sequence $\{B_n\}$ and $\{z_o\} \equiv \{(x_o, y_o)\} \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ denote the mild solution of equations (2.1)-(2.2) corresponding to B_o. To prove continuity we must show that $z_n \xrightarrow{s} z_o$ in $B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$. Subtracting the expressions (4.1) and (4.2) corresponding to $\{B_o, x_o, y_o\}$ from the same corresponding to $\{B_n, x_n, y_n\}$ we have, for all $t \in I$,

$$\begin{aligned} x_o(t) - x_n(t) &= \int_0^t S(t-s)[F(s,x_o(s)) - F(s,x_n(s))]ds \\ &+ \int_0^t S(t-s)(B_o(s)y_o(s) - B_n(s)y_n(s))ds \\ &+ \int_0^t S(t-s)[G(s,x_o(s)) - G(s,x_n(s))]dW(s), \end{aligned} \tag{5.1}$$

$$\begin{aligned} y_o(t) - y_n(t) &= \int_0^t S_0(t-s)[F_0(s,x_o(s)) - F_0(s,x_n(s))]ds \\ &+ \int_0^t S_0(t-s)[G_0(s,y_o(s)) - G_0(s,y_n(s))]dW_0(s). \end{aligned} \tag{5.2}$$

Considering the second term of the expression on the righthand side of equation (5.1) and denoting it by $(T2)$ we can rearrange it as

$$\begin{aligned} (T2)(t) &\equiv \int_0^t S(t-s)[B_o(s) - B_n(s)]y_o(s)ds \\ &+ \int_0^t S(t-s)B_n(s)[y_o(s) - y_n(s)]ds \\ &\equiv e_n(t) + \int_0^t S(t-s)B_n(s)[y_o(s) - y_n(s)]ds, t \in I. \end{aligned} \tag{5.3}$$

Considering the third term and denoting it by

$$(T3)(t) \equiv \int_0^t S(t-s)[G(s,x_o(s)) - G(s,x_n(s))]dW(s), t \in I,$$

we note that

$$\mathbf{E} \parallel (T3)(t) \parallel_X^2 = \mathbf{E} \int_0^t \parallel S(t-s)[G(s,x_o(s)) - G(s,x_n(s))] \parallel^2_{\gamma(H,X)} ds, \tag{5.4}$$

where $\parallel \cdot \parallel_{\gamma(H,X)}$ denotes the γ-Radonifying norm. Using the assumptions (A1), (A2)(ii) and (A3)(ii) and the equations (5.1), (5.3) and (5.4) and computing the expected value of the square of the X norm of $x_o(t) - x_n(t)$, we find that

$$\begin{aligned} &\mathbf{E} \parallel x_o(t) - x_n(t) \parallel_X^2 \\ &\le 4\big\{(MC_2)^2 t + (MC_4)^2\big\} \int_0^t \mathbf{E} \parallel x_o(t) - x_n(t) \parallel_X^2 ds \end{aligned} \tag{5.5}$$

$$+4(Mb)^2 t \int_0^t \mathbf{E} \parallel y_o(s) - y_n(s) \parallel_Y^2 ds + 4\mathbf{E} \parallel e_n(t) \parallel_Y^2, t \in I.$$

Similarly, considering the expression given by (5.2) and using the assumptions (A4), (A5(ii) and (A6)(ii) and computing the expected value of the square of the Y norm, one can easily verify that

$$\mathbf{E} \parallel y_0(t) - y_n(t) \parallel_Y^2 \leq 2(M_0 C_6)^2 t \int_0^t \mathbf{E} \parallel x_0(s) - x_n(s) \parallel_X^2 ds \tag{5.6}$$

$$+2(M_0 C_8)^2 \int_0^t \mathbf{E} \parallel y_0(s) - y_n(s) \parallel_Y^2 ds, t \in I.$$

Defining $\xi_n(t) \equiv \mathbf{E} \parallel x_o(t) - x_n(t) \parallel_X^2$ and $\eta_n(t) \equiv \mathbf{E} \parallel y_0(t) - y_n(t) \parallel_Y^2$ for $t \in I$, we can rewrite the inequalities (5.5)-(5.6) as follows:

$$\xi_n(t) \leq C_9 \int_0^t \xi_n(s) ds + C_{10} \int_0^t \eta_n(s) ds + 4\mathbf{E} \parallel e_n(t) \parallel_X^2 \tag{5.7}$$

$$\eta_n(t) \leq C_{11} \int_0^t \xi_n(s) ds + C_{12} \int_0^t \eta_n(s) ds, t \in I, \tag{5.8}$$

where the constants $C_9, C_{10}, C_{11}, C_{12}$ are dependent on T and the basic parameters $\{M, C_2, C_4\}$ and $\{M_0, C_6, C_8\}$ appearing in the assumptions (A1)-(A6). Defining $\zeta_n(t) = \xi_n(t) + \eta_n(t), t \in I$, and summing the above inequalities we obtain

$$\zeta_n(t) \leq \hat{C} \int_0^t \zeta_n(s)\, ds + 4\mathbf{E} \parallel e_n(t) \parallel_X^2, t \in I, \tag{5.9}$$

where $\hat{C} >$ is a constant dependent on the parameters $\{C_9, C_{10}, C_{11}, C_{12}\}$. It follows from Gronwall inequality applied to the above expression that

$$\zeta_n(t) \leq 4\mathbf{E} \parallel e_n(t) \parallel_X^2 + 4\hat{C} \exp(\hat{C}T) \int_0^t \mathbf{E} \parallel e_n(s) \parallel_X^2 ds, t \in I. \tag{5.10}$$

Recall that e_n is given by

$$e_n(t) \equiv \int_0^t S(t-s)[B_o(s) - B_n(s)] y_o(s) ds, t \in I. \tag{5.11}$$

Computing the square of the X norm of $e_n(t)$ we obtain the following inequality,

$$\parallel e_n(t) \parallel_X^2 \leq M^2 t \int_0^t \parallel [B_o(s) - B_n(s)] y_o(s) \parallel_X^2 ds \tag{5.12}$$

$$\leq M^2 T \int_0^T \parallel [B_o(s) - B_n(s)] y_o(s) \parallel^2 ds \; P - a.s.$$

Since $B_n \xrightarrow{\tau_{T\pi}} B_o$ (in Tychonoff product topology) and $y_o \in B_\infty^a(I, L_2(\Omega, Y))$, it is clear that the integrand converges P-a.s to zero for almost all $t \in I$. By our assumption, the elements of $\mathcal{B}_{ad}$ are uniformly norm bounded by a positive number b. Thus the integrand is dominated by the following integrable process

$$\parallel [B_o(t) - B_n(t)] y_o(t) \parallel^2 \leq 4b^2 \parallel y_o(t) \parallel_Y^2 .$$

Hence it follows from Lebesgue dominated convergence theorem that

$$\lim_{n\to\infty} \mathbf{E} \int_0^T \parallel [B_o(s) - B_n(s)]y_o(s) \parallel^2 ds \ = 0.$$

Thus we conclude that $\sup_{t\in I} \mathbf{E} \parallel e_n(t) \parallel_X^2 \longrightarrow 0$ as $n \to \infty$, and hence it follows from the inequality (5.10) that

$$\lim_{n\to\infty} \sup\{\zeta_n(t), t \in I\} = 0.$$

This proves that both ξ_n and η_n converges to zero uniformly on I and hence $z_n \stackrel{s}{\longrightarrow} z_o$ in $B_\infty^a(I, L_2(\Omega, \mathcal{Z}))$. This proves the continuity of the map $B \longrightarrow z$ with respect to the topologies stated. □

Using the above result we can prove the existence of optimal policy (an output feedback operator valued function). This is presented in the following theorem.

Theorem 5.3. *Consider the system (2.1)-(2.2) with the objective (cost) functional (2.3) and admissible output feedback control policies* $\mathcal{B}_{ad}$ *and suppose the assumptions of Theorem 5.2 hold. Further, suppose the cost integrands* $\{\ell, \Phi\}$ *are real valued Borel measurable functions on* $I \times X \times Y$ *and* $X \times Y$ *respectively and that for all* $t \in I$*, they are lower semicontinuous on* $X \times Y$ *satisfying the following growth conditions:*

$$|\ell(t, x, y)| \leq \alpha(t) + \beta\{\parallel x \parallel_X^2 + \parallel y \parallel_Y^2\}, |\Phi(x, y)| \leq \gamma\{\parallel x \parallel_X^2 + \parallel y \parallel_Y^2\}$$

for certain $\alpha \in L_1^+(I)$ *and real numbers* $\beta, \gamma \geq 0$*. Then, there exists a* $B_o \in \mathcal{B}_{ad}$ *at which* $J(B)$ *attains its minimum.*

Proof. For convenience, we write $J(B) \equiv J_1(B) + J_2(B)$ where

$$J_1(B) \equiv \mathbf{E} \int_I \ell(t, x, y) \ dt, \ J_2(B) \equiv \mathbf{E}\Phi(x(T), y(T)).$$

We show that these functionals are lower semicontinuous with respect to the Tychonoff product topology. Suppose $B_n \stackrel{\tau_{T\pi}}{\longrightarrow} B_o$, and let $z_n \equiv (x_n, y_n)$ denote the (mild) solution of the system (2.1)-(2.2) corresponding to the sequence B_n and $z_o \equiv (x_o, y_o)$ denote the solution corresponding to B_o. Then it follows from Theorem 5.2 that z_n converges to $z_o \equiv (x_o, y_o)$ in the norm topology of $B_\infty^a(I, L_2(\Omega, \mathcal{Z}))$. Since ℓ and Φ are lower semicontinuous on $X \times Y$, we have, for all $t \in I$,

$$\ell(t, x_o(t), y_o(t)) \leq \varliminf_{n\to\infty} \ell(t, x_n(t), y_n(t)) \ P - a.s \tag{5.13}$$

and

$$\Phi(x_o(T), y_o(T)) \leq \varliminf_{n\to\infty} \Phi(x_n(T), y_n(T)) \ P - a.s. \tag{5.14}$$

It follows from Remark 5.1 that the solution set $\mathcal{S}$ is a bounded subset of $B_\infty^a(I, L_2(\Omega, \mathcal{Z}))$. Consequently, it follows from our assumption on the quadratic growth property of the integrands $\{\ell, \Phi\}$ that they are bounded from below by integrable process. Thus it follows from generalized Fatou's Lemma that

$$\mathbf{E} \int_I \varliminf \ell(t, x_n(t), y_n(t)) dt \leq \varliminf \mathbf{E} \int_I \ell(t, x_n(t), y_n(t)) dt, \tag{5.15}$$

and hence it follows from (5.13) and (5.15) that $J_1(B_o) \leq \underline{\lim}\, J_1(B_n)$. It follows from similar argument that $J_2(B_o) \leq \underline{\lim}\, J_2(B_n)$. Thus the sum $J(B) = J_1(B) + J_2(B)$ is also lower semicontinuous with respect to the Tychonoff product topology $\tau_{\tau\pi}$. Since $\mathcal{B}_{ad}$ is compact in this topology, J attains its minimum on it. This proves the existence of an optimal policy. □

Remark 5.4. Since the map $B \longrightarrow J(B)$ is rarely convex, we cannot claim uniqueness of the optimal policy B_o. However, it is easy to verify that the set of optimals

$$\mathcal{O}p \equiv \{B \in \mathcal{B}_{ad} : J(B) = J(B_o)\}$$

is a closed subset of $\mathcal{B}_{ad}$ and hence, being a closed subset of a compact set, it is compact.

Remark 5.5. Considering the admissible set of (output feedback) control operators $\mathcal{B}_{ad} = B_0(I, \Lambda)$, with $\Lambda \subset \mathcal{L}_{so}(Y, X)$, we note that we have only admitted linear operator valued functions. We believe that this can be extended to a large class of nonlinear operators contained in X^Y. We leave it for future work.

6. Attainable set of measures and related control

It is interesting, both theoretically and practically, to study the properties of measures induced by the solutions of the stochastic evolution equations (2.1)-(2.2). We have seen that for each $B \in \mathcal{B}_{ad}$, the solution process $z(B) \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$. Let $\mathcal{B}(\mathcal{Z})$ denote the class of Borel sets in the product space $\mathcal{Z} \equiv X \times Y$. For any $B \in \mathcal{B}_{ad}$ and $t \in I$, let

$$m^B_t(\cdot) \equiv Pr\{z(B)(t) \in \cdot\}$$

denote the probability measure induced by the random element $z(B)(t)$ on $\mathcal{B}(\mathcal{Z})$. We are more interested in the probability measures induced by the state process $\{x(B)(t), t \in I\}$, not the observable process. This is given by the marginal

$$\mu^B_t(\Gamma) \equiv m^B_t(\Gamma \times Y), \Gamma \in \mathcal{B}(X).$$

For each $t \in I$, the attainable set of measures induced by the state process $\{x(B), B \in \mathcal{B}_{ad}\}$ is then given by

$$\mathcal{A}(t) \equiv \{\mu^B_t : B \in \mathcal{B}_{ad}\}$$

Theorem 6.1. *Suppose the assumptions of Theorem 5.2 hold. Then, for each $t \in I$, the attainable set of measures $\mathcal{A}(t)$ is a weakly compact subset of the space of regular Borel probability $\mathcal{M}_0(X)$.*

Proof. Let $(\mathcal{D}, \geq)$ be a directed set and $\mu^\alpha \in \mathcal{A}(t), \alpha \in \mathcal{D}$, be a net. Then by definition there exists a net $B_\alpha \in \mathcal{B}_{ad}$ so that $\mu^\alpha = \mu^{B_\alpha}_t$. Since $\mathcal{B}_{ad}$ is compact in the Tychonoff product topology $\tau_{\tau\pi}$, there exists a subnet, relabeled as the original net, and a $B_o \in \mathcal{B}_{ad}$ such that $B_\alpha \xrightarrow{\tau_{\tau\pi}} B_o$. Let $z_\alpha \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ denote the mild solution of the pair of evolution equations (2.1)-(2.2) corresponding to B_α and $z_o \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ the mild solution corresponding to B_o. Let x_α and x_o denote the projections of z_α and z_o respectively to their X components. That is, projection of $B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ to $B^a_\infty(I, L_2(\Omega, X))$. By Theorem 5.2, along a subnet if necessary, $z_\alpha \xrightarrow{s} z_o$ in $B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$. Thus $x_\alpha \xrightarrow{s} x_o$ in $B^a_\infty(I, L_2(\Omega, X))$ and hence, for each $t \in I$, $x_\alpha(t) \xrightarrow{s} x_o(t)$ in $L_2(\Omega, X)$. It follows from classical measure theory and probability theory that convergence in the mean implies convergence in measure which in turn implies convergence in distribution. It is known that convergence in distribution is equivalent to weak convergence of the corresponding measures. Thus $\mu_t^\alpha \xrightarrow{w} \mu_t^o$, that is, for each $\varphi \in BC(X)$,

$$\int_X \varphi(x)\mu_t^\alpha(dx) \longrightarrow \int_X \varphi(x)\mu_t^o(dx).$$

This proves that $\mu_t^o \in \mathcal{A}(t)$ and hence the attainable set is a weakly compact subset of the space of probability measures $\mathcal{M}_0(X)$. This completes the proof. □

Here we present some interesting applications. Suppose a desired (probability) measure $\nu \in \mathcal{M}_0(X)$ is given. The problem is to find a output feedback control law $B \in \mathcal{B}_{ad}$ so that, at the terminal time T, we have $\mu_T^B = \nu$. In order that this problem has a solution, it is necessary that $\nu \in \mathcal{A}(T)$. If this is true for every given $\nu \in \mathcal{M}_0(X)$, the system is globally controllable and, in that case, $\mathcal{A}(t)$ is weakly dense in $\mathcal{M}_0(X)$. If this is not satisfied, we can still try to find one that comes close to the desired ν. This requires a metric on the space of probability measures and it is well known that Prokhorov metric ρ is one that serves the purpose. For any $G \in \mathcal{B}(X)$, define the ε-neighbourhood of G by $G^\varepsilon \equiv \{x \in X : d(x, G) < \varepsilon\}$ where d is the metric induced by the norm topology of X. Recall that the Prokhorov metric between any two elements $\mu_1, \mu_2 \in \mathcal{M}_0(X)$ is given by

$$\rho(\mu_1, \mu_2) \equiv \inf\Big\{\varepsilon > 0 : \mu_1(G) \leq \mu_2(G^\varepsilon) + \varepsilon, \\ \text{and } \mu_2(G) \leq \mu_1(G^\varepsilon) + \varepsilon, G \in \mathcal{B}(X)\Big\}.$$

It is known that the metric topology on $\mathcal{M}_0(X)$ induced by the Prokhorov metric ρ is equivalent to the weak topology if and only if X is separable. Thus assuming separability of X we can formulate the problem stated above as follows: Find $B \in \mathcal{B}_{ad}$ that minimizes the following (cost) functional

$$J_1(B) \equiv \rho(\nu, \mu_T^B), B \in \mathcal{B}_{ad}.$$

Clearly this is equivalent to minimizing the functional $f(\mu)$ given by

$$f(\mu) \equiv \rho(\nu, \mu), \mu \in \mathcal{A}(T).$$

Corollary 6.2. *Consider the cost functional J_1 and suppose the assumptions of Theorem 6.1 hold and that X is separable. Then there exists an optimal feedback law $B_o \in \mathcal{B}_{ad}$ at which $J_1(B)$ attains its minimum.*

Proof. Under the separability assumption of the state space X, the metric topology on $\mathcal{M}_0(X)$ induced by the Prokhorov metric ρ, is equivalent to the weak topology. Hence, if μ^n converges weakly to μ^o, we have $\rho(\nu, \mu^n) \longrightarrow \rho((\nu, \mu^o)$. Thus the functional

$$\mu \longrightarrow f(\mu) \equiv \rho(\nu, \mu)$$

is weakly continuous on $\mathcal{M}_0(X)$. By Theorem 6.1, $\mathcal{A}(T)$ is weakly compact. Thus the functional $f(\mu)$ attains its minimum on $\mathcal{A}(T)$. Hence there exists a $B_o \in \mathcal{B}_{ad}$ at which J attains its minimum. This proves the Corollary. □

Another problem of practical interest is tracking a moving target. Let $\mathcal{K}(t), t \in I$, be a set valued function with closed values in the state space X. The problem is to find a control law $B \in \mathcal{B}_{ad}$ so that the functional

$$J_2(B) \equiv \int_0^T \mu_t^B(\mathcal{K}(t))dt$$

is maximized.

Corollary 6.3. *Consider the payoff functional J_2 and suppose the assumptions of Theorem 6.1 hold. Then there exists a feedback control law B_* at which J_2 attains its maximum.*

Proof. Let $\{B_n, B_o\} \in \mathcal{B}_{ad}$ be a sequence and $\{\mu^n, \mu^o\}$ the corresponding sequence of measure valued functions. It follows from Theorem 6.1, that for each $t \in I$, along a subsequence if necessary, $\mu_t^n \xrightarrow{w} \mu_t^o$. Since, for each $t \in I$, the set valued map $\mathcal{K}(t)$ has closed values, it follows from well known properties equivalent to weak convergence of probability measures that $\overline{\lim}\, \mu_t^n(\mathcal{K}(t)) \leq \mu_t^o(\mathcal{K}(t))$. As this holds for each $t \in I$, we conclude that

$$\overline{\lim}\, J_2 B_n) = \overline{\lim} \int_I \mu_t^n(\mathcal{K}(t))dt \leq \int_I \overline{\lim}\, \mu_t^n(\mathcal{K}(t))dt \leq \int_I \mu_t^o(\mathcal{K}(t))dt \equiv J_2(B_o).$$

This shows that J_2 is upper semicontinuous with respect to the Tychonoff product topology $\tau_{\tau\pi}$. Since $\mathcal{B}_{ad}$ is compact with respect this topology, J_2 attains its maximum on $\mathcal{B}_{ad}$. Thus there exists an optimal feedback control law in $\mathcal{B}_{ad}$. This completes the proof. □

Another interesting problem is related to obstacle avoidance. Let D be an open set in X. The problem is to find a control law that steers the system in such a way that it minimizes the contact probability with the obstacle D. One can approximately formulate this problem as an optimization problem. Find a $B \in \mathcal{B}_{ad}$ that minimizes the functional

$$J_3(B) \equiv \int_0^T \mu_t^B(D)\lambda(dt),$$

where λ is any finite positive Borel measure.

Corollary 6.4. *Consider the cost functional J_3 and suppose the assumptions of Theorem 6.1 hold. Then there exists a feedback control law B_* at which J_3 attains its minimum.*

Proof. Since D is an open set, following similar steps as in the proof of Corollary 6.3, one can verify that the functional $B \longrightarrow J_3(B)$ is lower semicontinuous in the Tychonoff product topology $\tau_{\tau\pi}$ on $\mathcal{B}_{ad}$. Thus compactness of $\mathcal{B}_{ad}$ in this topology implies the existence of a $B_* \in \mathcal{B}_{ad}$ at which the functional J_3 attains its minimum. This completes proof. □

Another related problem is concerned with the avoidance of multiple obstacles, different obstacles at different occasions during the time interval I. Let $\{D_i, i = 1, 2, \ldots, m\}$ be any family of disjoint open sets in the state space X and $\{t_i \in I, i =$

$1, 2, \ldots, m\}$ a set of specified instants of time. Let $\Psi : R_+^m \longrightarrow R$. The problem is to find a control law B that minimizes the functional

$$J_4(B) \equiv \Psi(\mu_{t_1}^B(D_1), \mu_{t_2}^B(D_2), \ldots, \mu_{t_m}^B(D_m)).$$

Theorem 6.5. *Consider the functional J_4 and suppose the assumptions of Theorem 6.1 hold and that $\Psi : R_+^m \longrightarrow R$ is a continuous and increasing function of its arguments and bounded away from $-\infty$. Then there exists a feedback control law B_* at which J_4 attains its minimum.*

Proof. We show that J_4 is lower semicontinuous. Let $\{B_n\}$ be any sequence in $\mathcal{B}_{ad}$ and suppose $B_n \xrightarrow{\tau_{\tau\pi}} B_o$. Let $\{\mu^n\}$ and μ^o denote the corresponding sequence of induced measure valued functions. It follows from Theorem 6.1 that for each $t \in I$, $\mu_t^n \xrightarrow{w} \mu_t^o$. Since each set D_i is open, it follows from properties equivalent to weak convergence that $\mu_{t_i}^o(D_i) \leq \underline{\lim}\, \mu_{t_i}^n(D_i)$ for each $i = \{1, 2 \ldots, m\}$. Clearly, it follows from the above inequality and the monotone increasing property of Ψ that

$$\Psi(\mu_{t_1}^o(D_1), \mu_{t_2}^o(D_2), \ldots, \mu_{t_m}^o(D_m)) \leq \Psi(\underline{\lim}\, \mu_{t_1}^n(D_1), \underline{\lim}\, \mu_{t_2}^n(D_2), \ldots, \underline{\lim}\, \mu_{t_m}^n(D_m)).$$

By definition of lower semicontinuity, for any $\varepsilon > 0$, there exists an integer n_ε such that, for all indices $i \in \{1, 2, \ldots, m\}$, $\underline{\lim}_{k\to\infty} \mu_{t_i}^k(D_i) \leq \mu_{t_i}^n(D_i) + \varepsilon$ for all $n > n_\varepsilon$. Since Ψ is an increasing function of its arguments, it follows from the previous inequality that

$$\Psi(\mu_{t_1}^o(D_1), \mu_{t_2}^o(D_2), \ldots, \mu_{t_m}^o(D_m)) \leq \Psi(\mu_{t_1}^n(D_1) + \varepsilon, \mu_{t_2}^n(D_2) + \varepsilon, \ldots, \mu_{t_m}^n(D_m) + \varepsilon),$$

for all $n > n_\varepsilon$. Hence, it is evident that

$$\Psi(\mu_{t_1}^o(D_1), \mu_{t_2}^o(D_2), \ldots, \mu_{t_m}^o(D_m)) \leq \underline{\lim}\, \Psi(\mu_{t_1}^n(D_1) + \varepsilon, \mu_{t_2}^n(D_2) + \varepsilon, \ldots, \mu_{t_m}^n(D_m) + \varepsilon).$$

Since the choice of $\varepsilon > 0$ is (otherwise) arbitrary, and Ψ is continuous, it follows from the above inequality that

$$\Psi(\mu_{t_1}^o(D_1), \mu_{t_2}^o(D_2), \ldots, \mu_{t_m}^o(D_m)) \leq \underline{\lim}\, \Psi(\mu_{t_1}^n(D_1), \mu_{t_2}^n(D_2), \ldots, \mu_{t_m}^n(D_m)).$$

Thus $J_4(B_o) \leq \underline{\lim}\, J_4(B_n)$. In other words, J_4 is lower semicontinuous on $\mathcal{B}_{ad}$ with respect to the $\tau_{\tau\pi}$ topology. Since $\mathcal{B}_{ad}$ is compact with respect to this topology and Ψ is bounded away from $-\infty$, J_4 attains its minimum on it and hence an optimal policy exists. This completes the proof. □

Another interesting problem is the escape time problem from a given set. Let $\mu_0 = \mathcal{L}(x_0)$ denote the probability law of the initial state and $C_0 \equiv \text{supp}(\mu_0)$ its support and suppose it is a closed bounded subset of X. Let $B_r \equiv B_r(X)$ be a closed ball of radius r in X containing the set C_0 in its interior. For any $B \in \mathcal{B}_{ad}$ and $\delta \in (0, 1)$, define

$$J_5(B) \equiv \inf\{t \geq 0 : \mu_t^B(B_r) < 1 - \delta\}.$$

If the underlying set is empty, we set $J_5(B) = T$. Otherwise, the problem is to find a $B \in \mathcal{B}_{ad}$ that maximizes the escape time. This is somewhat equivalent to choosing feedback control law to prevent finite time blowup (or equivalently increasing the life span of solution residing in the ball B_r.)

Theorem 6.6. *Consider the objective functional J_5 and suppose the assumptions of Theorem 6.1 hold. Then there exists a feedback control law $B_* \in \mathcal{B}_{ad}$ at which J_5 attains its maximum.*

Proof. It suffices to prove that $B \longrightarrow J_5(B)$ is upper semicontinuous with respect to the Tychonoff product topology $\tau_{\tau\pi}$. Let $B_n \in \mathcal{B}_{ad}$ and suppose B_n converges to B_o in the Tychonoff product topology $\tau_{\tau\pi}$. Let $\{\mu^n\}, \mu^o\}$ denote the corresponding induced (probability) measure valued functions. Since B_r is a closed ball, we have $\overline{\lim}\, \mu_t^n(B_r) \leq \mu_t^o(B_r)$. Thus it follows from a moment's reflection that

$$\{t \geq 0 : \overline{\lim}\, \mu_t^n(B_r) < 1-\delta\} \supset \{t \geq 0 : \mu_t^o(B_r) < 1-\delta\}.$$

Further, it is not difficult to verify that there exists an integer $n_o \in N$ such that

$$\{t \geq 0 : \overline{\lim}\, \mu_t^n(B_r) < 1-\delta\} \subset \{t \geq 0 : \mu_t^{n_o+k}(B_r) < 1-\delta\}$$

for all $k \in N$. From these inclusions we obtain the following inequality

$$\inf\{t \geq 0 : \mu_t^o(B_r) < 1-\delta\} \geq \inf\{t \geq 0 : \mu_t^{n_o+k}(B_r) < 1-\delta\}$$

for all $k \in N$. By definition of J_5, this is equivalent to $J_5(B_{n_o+k}) \leq J_5(B_o)$ for all $k \in N$. Hence $\overline{\lim}_k J_5(B_{n_o+k}) \leq J_5(B_o)$ proving upper semicontinuity of J_5. Since $\mathcal{B}_{ad}$ is compact in $\tau_{\tau\pi}$ topology, upper semicontinuity of J_5 in the same topology implies the existence of a control law $B_* \in \mathcal{B}_{ad}$ at which J_5 attains its maximum. This proves the existence. □

7. Necessary conditions of optimality

In this section we present necessary conditions of optimality for the control problem related to equations (2.1)-(2.3) whereby one can construct the optimal policy. We use the concept of Weak Radon Nikodym Property (WRNP). Let $\sigma(I)$ denote the sigma algebra of Lebesgue measurable subsets of the interval I and $\lambda(dt) \equiv dt$ denote the classical Lebesgue measure on $\sigma(I)$. A Banach space E is said to satisfy the weak Radon Nikodym property if, for every finitely additive measure μ defined on $\sigma(I)$ with values in E, there exists a Henstock-Kurzweil-Pettis integrable function g such that

$$\mu(D) = HKP \int_D g(t)\ \lambda(dt)$$

for every $D \in \sigma(I)$. For detailed characterization of RNP and WRNP see [9,12,14].

Theorem 7.1. *Consider the system (2.1)-(2.2) with the objective (cost) functional (2.3) and admissible feedback policies $\mathcal{B}_{ad} \equiv B_0(I,\Lambda)$ with Λ assumed to be a compact convex subset of the locally convex space $\mathcal{L}_{so}(Y,X)$. Suppose the assumptions of Theorem 5.3 hold and further, $\{F, G, F_0, G_0, \ell, \Phi\}$ are all once continuously Gâteaux differentiable with respect to the state and observation variables $\{x, y\}$ satisfying, along the trajectories, the properties (P1)-(P2):*

(P1): *$\{DF, DG, DF_0, DG_0\}$ are strongly measurable bounded operator valued functions,*

(P2): $\ell_x \in L_1(I, L_2(\Omega, X^*)), \ell_y \in L_1(I, L_2(\Omega, Y^*)), \Phi_x \in L_2(\Omega, X^*), \Phi_y \in L_2(\Omega, Y^*)$.

Let $B^o \in \mathcal{B}_{ad}$ *and* $z^o \equiv (x^o, y^o) \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ *the corresponding mild solution of the system (2.1)-(2.2). Then, in order for* $B^o \in \mathcal{B}_{ad}$ *to be optimal, it is necessary that there exists* $\psi \equiv (\psi_1, \psi_2) \in B^a_\infty(I, L_2(\Omega, X^*)) \times B^a_\infty(I, L_2(\Omega, Y^*)) \equiv B^a_\infty(I, L_2(\Omega, \mathcal{Z}^*))$ *such that the triple* $\{B^o, z^o, \psi\}$ *satisfies the inequality (7.1), the system equations (2.1)-(2.2) corresponding to* B^o*, and the adjoint evolution equations (7.2)-(7.3) as presented below:*

$$\mathbf{E}\int_0^T < (B - B^o)y^o, \psi_1 >_{X,X^*} dt \geq 0, \ \forall \ B \in \mathcal{B}_{ad}, \tag{7.1}$$

$$\begin{aligned} -d\psi_1 &= A^*\psi_1 dt + DF^*(t, x^o)\psi_1 dt + Q^*(t, x^o)\psi_1 dt + DF_0^*(t, x^o)\psi_2 dt \\ &\quad + \ell_x(t, x^o, y^o)dt + DG^*(t, x^o; \psi_1)dW, \\ \psi_1(T) &= \Phi_x(x^o(T), y^o(T)), \end{aligned} \tag{7.2}$$

$$\begin{aligned} -d\psi_2 &= A_0^*\psi_2 dt + (B^o)^*\psi_1 dt + Q_0^*(t, y^o)\psi_2 dt + \ell_y(t, x^o, y^o)dt \\ &\quad + DG_0^*(t, y^o; \psi_2)dW_0, \\ \psi_2(T) &= \Phi_y(x^o(T), y^o(T)), \end{aligned} \tag{7.3}$$

where the operators $\{Q, Q_0\}$ *are identified in the course of proof.*

Proof. By our assumption, Λ is a closed convex subset of $\mathcal{L}_{so}(Y, X)$ and hence the set of admissible control operator $\mathcal{B}_{ad}$ is also closed and convex. Thus, in order for $B^o \in \mathcal{B}_{ad}$ to be optimal, it is necessary that

$$J(B^o) \leq J(B^o + \varepsilon(B - B^o)) \ \forall \ B \in \mathcal{B}_{ad} \text{ and } \varepsilon \in [0, 1].$$

Hence, the Gâteaux differential dJ of J at B^o in the direction $(B - B^o)$ must satisfy the following inequality,

$$dJ(B^o, B - B^o) \geq 0 \ \forall \ B \in \mathcal{B}_{ad}. \tag{7.4}$$

Let $z^o = (x^o, y^o) \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ denote the mild solutions of the system (2.1)-(2.2) corresponding to $B^o \in \mathcal{B}_{ad}$ and $z^\varepsilon \equiv (x^\varepsilon, y^\varepsilon) \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ denote the mild solution corresponding to $B^\varepsilon \equiv B^o + \varepsilon(B - B^o)$. Clearly, these are the solutions of the integral equations (4.1)-(4.2). Define $v \equiv (v_1, v_2)$ with $v_1 = \lim_{\varepsilon \downarrow 0}(1/\varepsilon)(x^\varepsilon - x^o)$ and $v_2 \equiv \lim_{\varepsilon \downarrow 0}(1/\varepsilon)(y^\varepsilon - y^o)$. It follows from continuity of the map, $B \longrightarrow z$ (Theorem 5.2), and continuous Gâteaux differentiability of $\{F, G, F_0, G_0\}$ (with respect to the state variables (x, y)) that the limit v is well defined and it belongs to $B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ and it is the solution of the following system of integral equations:

$$\begin{aligned} v_1(t) &= \int_0^t S(t-s)DF(s, x^o(s); v_1(s))ds + \int_0^t S(t-s)B^o(s)v_2(s)ds \\ &\quad + \int_0^t S(t-s)(B(s) - B^o(s))y^o(s)ds \\ &\quad + \int_0^t S(t-s)DG(s, x^o(s); v_1(s))dW(s), t \in I, \end{aligned} \tag{7.5}$$

$$v_2(t) = \int_0^t S_0(t-s)DF_0(s, x^o(s); v_1(s))ds \tag{7.6}$$

$$+\int_0^t S_0(t-s)DG_0(s,y^o(s);v_2(s))dW_0(s), t\in I,$$

where $DF, DG; DF_0, DG_0$ denote the Gâteaux derivatives of $\{F, G, F_0, G_0\}$ in the directions as indicated. In other words $v \equiv (v_1, v_2) \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ is the mild solution of the following system of linear stochastic differential equations on the Banach space $\mathcal{Z} \equiv X \times Y$,

$$\begin{aligned} dv_1 &= Av_1dt + DF(t,x^o(t);v_1)dt + B^o(t)v_2dt \\ &\quad +(B-B^o)y^odt + DG(t,x^o(t);v_1)dW, v_1(0)=0, \end{aligned} \tag{7.7}$$

$$dv_2 = A_0v_2dt + DF_0(t,x^o(t);v_1)dt + DG_0(t,y^o(t);v_2)dW_0, v_2(0)=0. \tag{7.8}$$

It is clear from the above equations that if $B(t) = B^o(t), t \in I$, the system (7.7)-(7.8) reduces to a pair of homogeneous linear stochastic differential equations with zero initial condition. Hence it has only the trivial solution $v(t) \equiv (v_1(t), v_2(t)) \equiv (0,0), t \in I$. For convenience of presentation we denote

$$DF(t,x^o(t);v_1) \equiv DF^o(t)v_1,\ DF_0(t,x^o(t),v_1) \equiv DF_0^o(t)v_1,$$
$$DG(t,x^o(t);v_1) \equiv DG^o(t,v_1), DG_0(t,y^o(t);v_2) \equiv DG_0^o(t,v_2)$$

and note that they are all linear in v_1 and v_2. Writing the equations (7.7)-(7.8) in the form of a system,

$$\begin{aligned} dv &= \begin{bmatrix} A & 0 \\ 0 & A_0 \end{bmatrix} vdt + \begin{bmatrix} DF^o & B^o \\ DF_0^o & 0 \end{bmatrix} vdt + \\ &\quad + \begin{bmatrix} (B-B^o)y^o \\ 0 \end{bmatrix} dt + \begin{bmatrix} DG^o & 0 \\ 0 & DG_0^o \end{bmatrix} \begin{bmatrix} dW \\ dW_0 \end{bmatrix}, \end{aligned} \tag{7.9}$$

with initial value $v(0) = 0$, one observes that the system (7.9) is subject to the input $((B-B^o)y^o, 0)'$. By our assumptions, the operators DF^o and DF_0^o are strongly measurable bounded operator valued functions with values in $\mathcal{L}(X)$ and $\mathcal{L}(X,Y)$ respectively. Similarly, the operators DG^o and DG_0^o are also operator valued functions taking values from $\mathcal{L}(X, \gamma(H,X))$ and $\mathcal{L}(Y, \gamma(H_0, Y))$ respectively. Note that equation (7.9) is a linear stochastic differential equation on the (UMD) Banach space $\mathcal{Z} \equiv X \times Y$, and that the solution $v \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$. It is also clear that the map

$$(B-B^o)y^o, 0)' \longrightarrow v$$

is a continuous linear map from $B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$ to $B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$. Considering the cost functional J and computing its Gâteaux derivative at B^o in the direction $(B - B^o)$, we obtain

$$\begin{aligned} & dJ(B^o; B-B^o) \\ = \quad & \mathbf{E}\int_0^T \big\{< \ell_x^o(t), v_1(t) >_{X^*,X} + < \ell_y^o(t), v_2(t) >_{Y^*,Y}\big\}dt \\ & + \mathbf{E}\big\{< \Phi_x^o(T), v_1(T) >_{X^*,X} + < \Phi_y^o(T), v_2(T) >_{Y^*,Y}\big\} \end{aligned} \tag{7.10}$$

where

$$\ell_x^o(t) = \ell_x(t,x^o(t),y^o(t)), \ell_y^o(t) = \ell_y(t,x^o(t),y^o(t)), t \in I,$$

$$\Phi_x^o(T) = \Phi_x(x^o(T), y^o(T)), \Phi_y^o(T) = \Phi_y(x^o(T), y^o(T)).$$

Denote the expression on the right hand side of equation (7.10) by

$$(7.11) \qquad L(v) \equiv \mathbf{E}\int_0^T \big\{< \ell_x^o(t), v_1(t) >_{X^*,X} + < \ell_y^o(t), v_2(t) >_{Y^*,Y}\big\}dt$$
$$+\mathbf{E}\big\{< \Phi_x^o(T), v_1(T) >_{X^*,X} + < \Phi_y^o(T), v_2(T) >_{Y^*,Y}\big\}.$$

Using the properties (P2) and Hölder inequality applied to (7.11), it is easy to verify that there exists a finite positive number C such that

$$|L(v)| \leq C \parallel v \parallel_{B_\infty^a(I,L_2(\Omega,\mathcal{Z}))} .$$

Hence $v \longrightarrow L(v)$ is a continuous linear functional on $B_\infty^a(I, L_2(\Omega, \mathcal{Z}))$. Thus the composition map

$$(7.12) \qquad ((B - B^o)y^o, 0)^{'} \longrightarrow v \longrightarrow L(v) \equiv \tilde{L}(((B - B^o)y^o, 0)^{'})$$

is a continuous linear functional on $B_\infty^a(I, L_2(\Omega, \mathcal{Z}))$. The topological dual of this space is given by the space of $\mathcal{F}_t$-adapted $L_2(\Omega, \mathcal{Z}^*)$-valued finitely additive vector measures which we denote by $M_{fa}^a(\mathcal{I}, L_2(\Omega, \mathcal{Z}^*))$ where $\mathcal{I}$ denotes the sigma algebra of Lebesgue measurable subsets of the interval I. In other words, for every continuous linear functional ℓ on the Banach space $B_\infty^a(I, L_2(\Omega, \mathcal{Z}))$, there exists a unique

$$\mu \in (B_\infty^a(I, L_2(\Omega, \mathcal{Z})))^* = M_{fa}^a(\mathcal{I}, L_2(\Omega, \mathcal{Z}^*))$$

such that $\ell(f)$ has the following representation

$$\ell(f) = \mathbf{E}\int_I < f(t), \mu(dt) > .$$

Thus it follows from continuity of the map (7.12) and the above duality pairing that there exists a $\mu \in M_{fa}^a(\mathcal{I}, L_2(\Omega, \mathcal{Z}^*))$ such that

$$L(v) = \tilde{L}(((B - B^o)y^o, 0)^{'}) = \mathbf{E}\int_I < (B - B^o)y^o, 0)^{'}, \mu(dt) >_{\mathcal{Z},\mathcal{Z}^*} .$$

Since, by assumption, both X and Y are UMD Banach spaces, they are reflexive and hence $\mathcal{Z} \equiv (X \times Y)$ is a reflexive Banach space and therefore it contains no isomorphic copy of ℓ_1. Thus it follows from a theorem due to Janicka [18], also reported in Musail and Ryll-Nardzewski [**?**, Theorem 6, p83][14] including its converse, that $\mathcal{Z}^*$ satisfies weak Radon-Nikodym property (WRNP). Hence it follows from a result due to Bongiorno-DiPazza-Musial [9, Theorem 4.5, p481], that there exists an $\mathcal{F}_t$-adapted Henstock-Kurzweil-Pettis integrable function $\psi = (\psi_1, \psi_2) \in L_{HKP}^a(I, L_2(\Omega, \mathcal{Z}^*))$ such that the above expression can be rewritten as

$$(7.13) \qquad L(v) = \tilde{L}(((B - B^o)y^o, 0)^{'}) = \mathbf{E}\int_I < (B - B^o)y^o, 0)^{'}, \mu(dt) >_{\mathcal{Z},\mathcal{Z}^*}$$
$$= \mathbf{E}\int_I < ((B - B^o)y^o, 0), \psi(t) >_{\mathcal{Z},\mathcal{Z}^*} dt$$
$$= \mathbf{E}\int_I \{< (B - B^o)y^o, \psi_1 >_{X,X^*} + < 0, \psi_2 >_{Y,Y^*}\}dt$$

$$= \mathbf{E} \int_I \{< (B - B^o)y^o, \psi_1 >_{X,X^*}\} dt.$$

It is well known that a reflexive Banach space E has RNP (Radon Nikodym Property). Therefore, for any finite measure space $(\Omega, \Sigma, \lambda)$ and any λ-continuous E valued countably additive vector measure m, there exists a Bochner integrable E valued function h such that, for every $S \in \Sigma$, $m(S) = \int_S h(s) d\lambda(s)$. Indefinite integrals in the sense of Bochner, as well as Pettis, are countably additive measures. For an excellent account on the subject see [12,13]. In our case the Z^* valued vector measure μ is only finitely additive and this is the primary reason for using Henstock-Kurzweil-petits integral. So far we have established the existence of $\psi = (\psi_1, \psi_2)$ giving an alternate expression for the functional $L(v)$ and hence the directional derivative of J. We must now develop a constructive procedure to determine ψ while eliminating the variational equations. We can do this using the above expression and the variational equations (7.7)-(7.8) or equivalently (7.9). For this we consider the real valued functional g defined on $\mathcal{Z} \times \mathcal{Z}^*$ and given by

$$g(v_1(t), v_2(t), \psi_1(t), \psi_2(t)) \\ =< v_1(t), \psi_1(t) >_{X,X^*} + < v_2(t), \psi_2(t) >_{Y,Y^*}, t \in I. \tag{7.14}$$

With respect to compatible duality pairings, the Itô differential rules remain valid also in UMD Banach spaces [10, 19]. Thus

$$\begin{aligned} dg &= d < v_1, \psi_1 > + d < v_2, \psi_2 > \\ &= \{< v_1, d\psi_1 > + < dv_1, \psi_1 > + << dv_1, d\psi_1 >>\} \\ &\quad + \{< v_2, d\psi_2 > + < dv_2, \psi_2 > + << dv_2, d\psi_2 >>\}, \end{aligned} \tag{7.15}$$

where $<< dv_1, d\psi_1 >>$ and $<< dv_2, d\psi_2 >>$ denote the quadratic variation terms. Integrating dg by parts and recalling that $v_1(0) = 0, v_2(0) = 0$, we obtain

$$\begin{aligned} \mathbf{E} \int_0^T dg &= \mathbf{E} \int_0^T \{d < v_1(t), \psi_1(t) > + d < v_2(t), \psi_2(t) >\} dt \\ &= \mathbf{E} < v_1(T), \psi_1(T) >_{X,X^*} + \mathbf{E} < v_2(T), \psi_2(T) >_{Y,Y^*}. \end{aligned} \tag{7.16}$$

Next, we consider the first term on the RHS of equation (7.15). For convenience of presentation, let us denote its integral by $R1$,

$$R1 \equiv \mathbf{E} \int_0^T < v_1, d\psi_1 > + < dv_1, \psi_1 > + << dv_1, d\psi_1 >> . \tag{7.17}$$

Using the variational equation (7.7) and formally substituting in the above expression and integrating by parts we arrive at the following expression

$$\begin{aligned} R1 = \mathbf{E} \int_0^T \Big\{ &< v_1, d\psi_1 + A^* \psi_1 dt + DF^*(t, x^o; \psi_1) dt > \\ &+ < v_1, DG^*(t, x^o; \psi_1) dW > \\ &+ < v_2, (B^o)^* \psi_1 dt > + < (B - B^o) y^o, \psi_1 dt > + << dv_1, d\psi_1 >> \Big\}. \end{aligned} \tag{7.18}$$

Similarly, we consider the second term on the RHS of equation (7.15) and denote its integral by $R2$ as follows:

$$R2 \equiv \mathbf{E}\int_0^T \Big\{ < v_2, d\psi_2 > + < dv_2, \psi_2 > + << dv_2, d\psi_2 >> \Big\}. \tag{7.19}$$

Using the variational equation (7.8) and again formally substituting in the above expression we arrive at the following identity

$$\begin{aligned} R2 = \mathbf{E}\int_0^T \Big\{ &< v_2, d\psi_2 + A_0^*\psi_2 dt + DG_0^*(t, y^o; \psi_2)dW_0 > \\ &+ < v_1, DF_0^*(t, x^o; \psi_2)dt > + << dv_2, d\psi_2 >> \Big\}. \end{aligned} \tag{7.20}$$

The formal substitution is justified by using the Yosida approximation [1] of the identity operators in X and Y generated by the unbounded operators A and A_0 respectively. Using the resolvents $R(n, A), n \in \rho(A)$, where $\rho(A)$ is the resolvent set of A and $R(n, A_0), n \in \rho(A_0)$, we obtain the approximations of the identity operators $I_n \equiv nR(n, A)$ on X and $I_{0,n} = nR(nA_0)$ on Y respectively. Clearly, $Range(I_n) \subset D(A)$ and $Range(I_{0,n}) \subset D(A_0)$ for all $n \in N$. Further, both I_n and $I_{0,n}$ converge in the strong operator topology to the identity operators $I \in \mathcal{L}(X)$ and $I_0 \in \mathcal{L}(Y)$ respectively. All the operators in the variational equations (7.7)-(7.8) except $\{A, A_0\}$ are then replaced by their Yosida approximations. This leads to a sequence of variational equations giving a corresponding sequence of solutions $\{v_{1,n}, v_{2,n}\}$ which take values in the domains of the operators A and A_0 respectively. Then these regularized equations are used to carry out all the intermediate operations, and then, finally letting $n \to \infty$, we reach the above expressions (R1) and (R2). Next, we consider the quadratic variation terms. Since, by our assumptions, G and G_0 are continuously Gâteaux differentiable, the corresponding directional derivatives are bounded strongly measurable operator valued functions with values in the space of γ-Radonifying operators as indicated below:

$$DG(t, x^o(t); v_1(t)) \in \gamma(H, X), \text{ and } DG_0(t, y^o(t); v_2(t)) \in \gamma(H_0, Y)$$

for almost all $t \in I$, P-a.s. Accordingly their duals belong to the corresponding dual spaces as displayed below:

$$DG^*(t, x^o(t); \psi_1(t)) \in \gamma(H, X^*) \text{ and } DG_0^*(t, y^o(t); \psi_2(t)) \in \gamma(H_0, Y^*)$$

for almost all $t \in I$, P-a.s. Therefore, considering the quadratic variation terms, recalling the independence of the two H and H_0 Brownian motions, and noting that $\{DG, DG_0\}$ and $\{DG^*, DG_0^*\}$ are linear in their third argument, we obtain from equations (7.7)-(7.8) and (7.18)-(7.20) the following bilinear forms:

$$\begin{aligned} &\mathbf{E}\int_0^T \{ << dv_1, d\psi_1 >> \} \\ &= \mathbf{E}\int_0^T \{<< DG(t, x^o; v_1)dW, (-1)DG^*(t, x^o; \psi_1)dW >>\} \\ &= \mathbf{E}\int_0^T dt < DG(t, x^o(t); v_1(t)), (-)DG^*(t, x^o(t); \psi_1(t)) >_{\gamma(H,X),\gamma(H,X^*)} \end{aligned} \tag{7.21}$$

$$\equiv \mathbf{E}\int_0^T (Q(t,x^o(t))v_1(t),\psi_1(t))_{X,X^*}dt,$$

and

(7.22)

$$\begin{aligned}
&\mathbf{E}\int_0^T \{<< dv_2, d\psi_2 >>\} \\
&\quad= \mathbf{E}\int_0^T \{<< DG_0(t,y^o;v_2)dW_0, (-1)DG_0^*(t,y^o;\psi_2)dW_0 >>\} \\
&\quad\equiv \mathbf{E}\int_0^T dt < DG_0(t,y^o(t);v_2(t)), (-)DG_0^*(t,y^o(t);\psi_2(t)) >_{\gamma(H_0,Y),\gamma(H_0,Y^*)} \\
&\quad\equiv \mathbf{E}\int_0^T (Q_0(t,y^o(t))v_2(t),\psi_2(t))_{Y,Y^*}dt.
\end{aligned}$$

By our assumption, both G and G_0 are continuously Gâteaux differentiable in their second argument. Hence the operators $Q(t) \equiv Q(t,x^o(t))$ and $Q_0(t) \equiv Q_0(t,y^o(t))$ are well defined strongly measurable operator valued functions on I taking values in $\mathcal{L}(X)$ and $\mathcal{L}(Y)$) respectively. Summing (7.18), (7.19), (7.21) and (7.22) and equating with (7.16) we obtain

$$\begin{aligned}
(7.23)\qquad &\mathbf{E} < v_1(T),\psi_1(T) >_{X,X^*} + \mathbf{E} < v_2(T),\psi_2(T) >_{Y,Y^*} \\
&= \mathbf{E}\int_0^T \big\langle v_1, d\psi_1 + A^*\psi_1 dt + DF^*(t,x^o;\psi_1)dt + DF_0^*(t,x^o;\psi_2)dt \\
&\quad + Q^*(t,x^o)\psi_1 dt + DG^*(t,x^o;\psi_1)dW\big\rangle \\
&\quad + \mathbf{E}\int_0^T \big\langle v_2, d\psi_2 + A_0^*\psi_2 dt + (B^o)^*\psi_1 dt \\
&\quad + Q_0^*(t,y^o)\psi_2 dt + DG_0^*(t,y^o;\psi_2)dW_0\big\rangle \\
&\quad + \mathbf{E}\int_0^T < (B-B^o)y^o,\psi_1 >_{X,X^*} dt.
\end{aligned}$$

Setting

$$\begin{aligned}
(7.24)\qquad & d\psi_1 + A^*\psi_1 dt + DF^*(t,x^o;\psi_1)dt + DF_0^*(t,x^o;\psi_2)dt \\
&\qquad + Q^*(t,x^o)\psi_1 dt + DG^*(t,x^o;\psi_1)dW = -\ell_x(t,x^o,y^o)dt,
\end{aligned}$$

$$(7.25)\qquad \psi_1(T) = \Phi_x^o(T) \equiv \Phi_x(x^o(T),y^o(T));$$

and

$$\begin{aligned}
(7.26)\qquad & d\psi_2 + A_0^*\psi_2 dt + (B^o)^*\psi_1 dt + Q_0^*(t,y^o)\psi_2 dt \\
&\qquad + DG_0^*(t,y^o;\psi_2)dW_0 = -\ell_y(t,x^o,y^o)dt,
\end{aligned}$$

$$(7.27)\qquad \psi_2(T) = \Phi_y^o(T) \equiv \Phi_y(x^o(T),y^o(T));$$

it follows from the expression (7.23) that the following identity holds,

$$\begin{aligned}
(7.28)\qquad & \mathbf{E} < v_1(T),\Phi_x^o(T) >_{X,X^*} + \mathbf{E} < v_2(T),\Phi_y^o(T) >_{Y,Y^*} \\
&\quad + \mathbf{E}\int_0^T < v_1(t),\ell_x(t,x^o,y^o) >_{X,X^*} dt
\end{aligned}$$

$$+\mathbf{E}\int_0^T < v_2(t), \ell_y(t,x^o,y^o) >_{Y,Y^*} dt$$
$$= \mathbf{E}\int_0^T < (B-B^o)y^o, \psi_1 >_{X,X^*} dt.$$

Note that the expression on the left hand side of the above equation coincides with the functional $L(v)$ given by equation (7.10). Hence it follows from (7.4) and the above expression that

$$\mathbf{E}\int_0^T < (B-B^o)y^o, \psi_1 >_{X,X^*} dt \geq 0, \forall\ B \in \mathcal{B}_{ad}. \tag{7.29}$$

This gives the necessary condition (7.1). The necessary conditions (7.2)-(7.3) follow from the equations (7.24)-(7.25) and (7.26)-(7.27). The mild solutions of the adjoint equations (7.24)-(7.25) and (7.26)-(7.27) are given, respectively, by the solutions of the following linear backward stochastic integral equations,

$$\begin{aligned}\psi_1(t) = S^*(T-t)\Phi_x^o(T) &+ \int_t^T S^*(s-t)DF^*(s,x^o(s);\psi_1(s))ds \\ &+ \int_t^T S^*(s-t)DF_0^*(s,x^o(s);\psi_2(s))ds \\ &+ \int_t^T S^*(s-t)Q^*(s,x^o(s))\psi_1(s)ds \\ &+ \int_t^T S^*(s-t)\ell_x(s,x^o(s),y^o(s))ds \\ &+ \int_t^T S^*(s-t)DG^*(s,x^o(s);\psi_1(s))dW(s), t \in I,\end{aligned} \tag{7.30}$$

$$\begin{aligned}\psi_2(t) = S_0^*(T-t)\Phi_y^o(T) &+ \int_t^T S_0^*(s-t)(B^o(s))^*\psi_1(s)ds \\ &+ \int_t^T S_0^*(s-t)Q_0^*(s,y^o(s))\psi_2(s))ds \\ &+ \int_t^T S_0^*(s-t)\ell_y(s,x^o(s),y^o(s))ds \\ &+ \int_t^T S_0^*(s-t)DG_0^*(s,y^o(s);\psi_2(s))dW_0(s), t \in I.\end{aligned} \tag{7.31}$$

Using the assumptions (P1) and (P2) and a standard approach based on successive Piccard approximation technique developed by Hu and Peng [**?**, Theorem 3.1]] for backward stochastic differential equations (BSDE), one can prove the existence and uniqueness of a solution $\psi = (\psi_1, \psi_2)$ of equations (7.30)-(7.31) in the Banach space $B_\infty^a(I, L_2(\Omega, \mathcal{Z}^*))$. This completes the proof. □

Remark 7.2. It is interesting to note that the solution ψ of the adjoint system (7.2)-(7.3) predicted by the WRNP (Weak Radon Nikodym property) is Henstock-Kurzweil-Pettis integrable, that is, $\psi \in L_{HKP}^a(I, L_2(\Omega, \mathcal{Z}^*)$. On the other hand, by application of Piccard approximation technique for backward stochastic differential equations (BSDE), we find that the adjoint system has a unique mild solution $\psi \in B^a_\infty(I, L_2(\Omega, Z^*))$ which is clearly much more regular.

Remark 7.3. Since the state x of system (2.1) is not accessible, the cost integrands $\{\ell, \Phi\}$ may not be directly dependent on x, though the cost functional J is indirectly dependent on state process. In this case the necessary conditions of optimality given by Theorem 7.1 simplify with no change in the inequality (7.1), and $\psi_1(T) = 0, \ell_x \equiv 0$ for the adjoint equation (7.2), and no change in equation (7.3).

Remark 7.4. In this section we have developed necessary conditions of optimality for control problem (2.3) known as the Bolza problem. For optimization problems involving measure valued functions as in section 6, we can develop similar necessary conditions of optimality. This requires the notion of measure solutions or equivalently evolution equations on the space of measures and their control as seen in Ahmed [4] and the references therein. The author believes that similar results can be proved on UMD Banach spaces.

8. Convergence theorem

Using Theorem 7.1, one can determine the optimal feedback operator. For this we construct a sequence of operators based on the necessary conditions and prove that the sequence converges monotonically to the optimal operator (more precisely the local optimal). Let E and F be a pair of real Banach spaces and let $\mathcal{L}_1(E, F)$ denote the space of nuclear operators. An operator $K \in \mathcal{L}_1(E, F)$ has the representation $K \equiv \sum e_i^* \otimes f_i$, where $\{e_i^*\} \in E^*$ and $\{f_i\} \in F$ are all linearly independent vectors satisfying the property $\lim_{i\to\infty} \parallel e_i^* \parallel_{E^*} = \lim_{i\to\infty} \parallel f_i \parallel_F = 0$, with the nuclear norm given by

$$\parallel K \parallel_{\mathcal{L}_1(E,F)} \equiv \sum_{i\geq 1} \parallel e_i^* \parallel_{E^*} \parallel f_i \parallel_F .$$

Another representation of K is given by $K = \sum \lambda_i \, e_i^* \otimes f_i$ where now the vectors $\{e_i^*, f_i\}$ can be chosen to be normalized in the sense that $e_i^* \in B_1(E^*), f_i \in B_1(F)$ and $\lambda \in \ell_1$. In this case $\parallel K \parallel_{\mathcal{L}_1(E,F)} = \sum |\lambda_i| \equiv Tr(K)$. It is well known that the topological dual of $\mathcal{L}_1(E, F)$ is given by the space of bounded linear operators from E^* to F^* denoted by $\mathcal{L}(E^*, F^*)$. The duality pairing between elements of $\mathcal{L}(E^*, F^*)$ with those of $\mathcal{L}_1(E, F)$ is given by

$$< L, K >_{\mathcal{L}(E^*,F^*),\mathcal{L}_1(E,F)} = \sum (Le_i^*, f_i)_{F^*,F}.$$

Clearly, $| < L, K > | \leq \parallel L \parallel_{\mathcal{L}(E^*,F^*)} \parallel K \parallel_{\mathcal{L}_1(E,F)}$. Using the tensor product notation we can rewrite the necessary condition (7.1) as follows:

$$J(B^o, B - B^o) = \mathbf{E} \int_0^T < (B - B^o)y^o, \psi_1 >_{X,X^*} dt \tag{8.1}$$

$$= \mathbf{E} \int_0^T < (B(t) - B^o(t)), N_o(t) >_{\mathcal{L}(Y,X),\mathcal{L}_1(Y^*,X^*)} dt$$

where $N_o(t) \equiv \{y^o(t) \otimes \psi_1(t)\}, t \in I$. Since X, Y are reflexive Banach spaces, the dual of $\mathcal{L}_1(Y^*, X^*)$ is given by $\mathcal{L}(Y, X)$. For the proof of convergence theorem we need the duality map denoted by $\triangle$. For each $N \in \mathcal{L}_1(Y^*, X^*)$, define the set

$$\triangle(N) \equiv \{L \in \mathcal{L}(Y, X) :< L, N >=\| L \|^2_{\mathcal{L}(Y,X)} =\| N \|^2_{\mathcal{L}_1(Y^*,X^*)}\}.$$

By Hahn-Banach theorem this is a nonempty closed convex set. With this preparation, we can now prove the following convergence theorem.

Theorem 8.1. *Suppose the assumptions of Theorem 7.1 hold and further assume that Y is a separable (reflexive) Banach space. Then there exists a number $m_0 \in R$ and a sequence $\{B_n\} \in \mathcal{B}_{ad}$ along which J is monotone decreasing and*

$$\lim_{n\to\infty} J(B_n) = m_0 > -\infty.$$

Proof. Since we want the feedback operator to be deterministic, we rewrite the expression (8.1) as follows

$$J(B^o, B - B^o) = \int_0^T < (B(t) - B^o(t)), \hat{N}_o(t) >_{\mathcal{L}(Y,X),\mathcal{L}_1(Y^*,X^*)} dt \tag{8.2}$$

where $\hat{N}_o(t) \equiv \mathbf{E}\{N_o(t)\} = \int_\Omega N_o(t, \omega)P(d\omega), t \in I$. This is justified by use of Fubini's theorem. Next, we choose an arbitrary $B_1 \in \mathcal{B}_{ad}$ and solve the pair of equations (2.1)-(2.2) in the mild sense and denote these mild solutions by $\{x_1, y_1\}$. Use these solutions in the adjoint pair of equations (7.2)-(7.3) by replacing the triple $\{B^o, x^o, y^o\}$ by $\{B_1, x_1, y_1\}$, and solve for the pair $\{\psi_{1,1}, \psi_{2,1}\}$ yielding the quintuple $\{B_1, x_1, y_1, \psi_{1,1}, \psi_{2,1}\}$. Using this quintuple, we construct N_1 given by

$$N_1(t) \equiv y_1(t) \otimes \psi_{1,1}(t), t \in I.$$

Since $y_1 \in B^a_\infty(I, L_2(\Omega, Y))$ and $\psi_{1,1} \in B^a_\infty(I, L_2(\Omega, X^*))$ it is clear that, for almost all $t \in I$ and P-a.s, $N_1(t) \in \mathcal{L}_1(Y^*, X^*)$ and that it is Bochner integrable in the sense that

$$\hat{N}_1(t) \equiv \mathbf{E}N_1(t) \equiv \int_\Omega N_1(t, \omega)P(d\omega).$$

Clearly, $\hat{N}_1$ is a strongly measurable function on I with values in the Banach space $\mathcal{L}_1(Y^*, X^*)$. Let $\mathcal{L}_{so}(Y, X)$ denote the space of bounded linear operators from Y to X endowed with the strong operator topology τ_{so}. This is a locally convex sequentially complete Hausdorff topological space. Since Y is assumed to be separable, this topology is metrizable with the metric

$$d(T, L) \equiv \sum_{n=1}^\infty (1/2^n) \min\{1, \| (T - L)\zeta_n \|_X\} \text{ for } T, L \in \mathcal{L}_{so}(Y, X),$$

where $\{\zeta_n\}$ is a dense subset of the closed unit ball $B_1(Y)$. We denote the completion of $\mathcal{L}_{so}(Y, X)$ with respect to the above metric by $\mathcal{L}_{sod}(Y, X)$. This is a complete metric space but not separable. Clearly, the strong operator topology is equivalent to the metric topology. Using the metric topology one can verify that the graph

$$\mathcal{G}r(\triangle) \equiv \{(L, N) \in \mathcal{L}(Y, X) \times \mathcal{L}_1(Y^*, X^*) : L \in \triangle(N)\}$$

of the multifunction $\triangle$ is closed with respect the norm topology of $\mathcal{L}_1(Y^*, X^*)$ and the metric topology of $\mathcal{L}_{sod}(Y, X)$. Thus the duality map

$$\triangle : \mathcal{L}_1(Y^*, X^*) \longrightarrow \mathcal{L}_{sod}(Y, X) \equiv (\mathcal{L}_{so}(Y, X), d)$$

is an upper semi-continuous multi function. Since $t \longrightarrow \hat{N}_1(t)$ is strongly measurable, the composition map $I \ni t \longrightarrow \hat{\triangle}_1(t) \equiv (\triangle \circ \hat{N}_1)(t) \equiv \triangle(\hat{N}_1(t))$ is a weakly measurable multi function in the sense that, for every open set $\mathcal{O} \subset \mathcal{L}_{sod}(Y, X)$, the set $\{t \in I : \hat{\triangle}_1(t) \cap \mathcal{O} \neq \emptyset\}$ is Lebesgue measurable. We have seen that the metric space $\mathcal{L}_{sod}(Y, X)$ is complete but not separable, and so, not a Polish or even a Souslin space. Thus, unfortunately, many measurable selection theorems requiring the target space to be separable [**?**, Theorem 1, p26]] are not applicable. Since $\mathcal{L}_{sod}(Y, X)$ with the metric topology is not σ-compact either, measurable selection theorems, requiring the target space to be σ-compact, are also not applicable. Under some mild assumptions on additivity and reducibility (with respect to the class of Lebesgue measurable subsets of the interval I), Graf presents, in an excellent survey paper [15, Theorem 3.5, p98] a general result on selection theorem which does not require separability. Based on this result [15, Theorem 3.5, p98], we can assert that the multifunction $\hat{\triangle}_1$ has measurable selections. Let L_1 be a measurable selection of $\hat{\triangle}_1$ in the sense that $L_1(t) \in \hat{\triangle}_1(t), t \in I$. Then define the operator valued function B_2 by setting

$$B_2(t) \equiv B_1(t) - \varepsilon L_1(t), t \in I,$$

for $\varepsilon > 0$ sufficiently small, so that $B_2 \in \mathcal{B}_{ad}$. Computing the cost functional J at B_2 we obtain $J(B_2) = J(B_1) + dJ(B_1, B_2 - B_1) + o(\varepsilon)$. Using the expression for the Gâteaux differential of J at B_1 in the direction $(B_2 - B_1)$ we arrive at the following expression,

$$\begin{aligned}
J(B_2) &= J(B_1) + \mathbf{E}\int_0^T < B_2 - B_1, N_1 >_{\mathcal{L}(Y,X),\mathcal{L}_1(Y^*,X^*)} dt + o(\varepsilon) \\
&= J(B_1) + \int_0^T < B_2 - B_1, \hat{N}_1(t) >_{\mathcal{L}(Y,X),\mathcal{L}_1(Y^*,X^*)} dt + o(\varepsilon) \\
&= J(B_1) - \varepsilon \int_0^T < L_1(t), \hat{N}_1(t) >_{\mathcal{L}(Y,X),\mathcal{L}_1(Y^*,X^*)} dt + o(\varepsilon) \\
&= J(B_1) - \varepsilon \parallel L_1 \parallel^2_{\mathcal{L}(Y,X)} + o(\varepsilon) = J(B_1) - \varepsilon \parallel \hat{N}_1 \parallel^2_{\mathcal{L}_1(Y^*,X^*)} + o(\varepsilon).
\end{aligned} \tag{8.3}$$

Thus, for $\varepsilon > 0$ sufficiently small, we have $J(B_2) < J(B_1)$. Next we use B_2 in equation (2.1) and solve the state equations (2.1)-(2.2) (in the mild sense) giving the pair $\{x_2, y_2\} \in B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$. Using this triple $\{B_2, x_2, y_2\}$ in place of $\{B^o, x^o, y^o\}$ in the adjoint system (7.2)-(7.3) and solving these equations we obtain the pair $\{\psi_{1,2}, \psi_{2,2}\}$. Thus we have the quintuple $\{B_2, x_2, y_2\}$, $\{\psi_{1,2}, \psi_{2,2}\}$. Using this solution we construct the nuclear operator valued function (process) $N_2(t) \equiv y_2(t) \otimes \psi_{1,2}(t), t \in I$, and $\hat{N}_2(t) \equiv \mathbf{E}(N_2(t)), t \in I$. Then using the duality map $\triangle$ we define the multi function $\hat{\triangle}_2$ given by the composition

$$\hat{\triangle}_2(t) \equiv (\triangle o \hat{N}_2)(t), t \in I.$$

Using the selection theorem as stated above, we choose a (measurable) selection $L_2(t) \in \hat{\triangle}_2(t), t \in I$, and define the operator valued function B_3 by $B_3(t) \equiv B_2(t) - \varepsilon L_2(t), t \in I$, so that $B_3 \in \mathcal{B}_{ad}$. Then, following similar steps as in the preceding case and computing the cost functional J at B_3, we find that

$$
\begin{aligned}
(8.4)\qquad J(B_3) &= J(B_2) + dJ(B_2, B_3 - B_2) + o(\varepsilon) \\
&= J(B_2) + \mathbf{E}\int_0^T < B_3 - B_2, N_2 >_{\mathcal{L}(Y,X),\mathcal{L}_1(Y^*,X^*)} dt + o(\varepsilon) \\
&= J(B_2) - \varepsilon \parallel L_2 \parallel^2_{\mathcal{L}(Y,X)} + o(\varepsilon) = J(B_2) - \varepsilon \parallel \hat{N}_2 \parallel^2_{\mathcal{L}_1(Y^*,X^*)} + o(\varepsilon).
\end{aligned}
$$

Thus, for $\varepsilon > 0$ sufficiently small, we have $J(B_3) < J(B_2) < J(B_1)$. Continuing this process we obtain a sequence $\{B_n\} \subset \mathcal{B}_{ad}$ satisfying

$$J(B_1) > J(B_2) > \cdots J(B_n) > J(B_{n+1} > \cdots .$$

It follows from Remark 5.1 that the solution set $\mathcal{S}$ is a bounded subset of $B^a_\infty(I, L_2(\Omega, \mathcal{Z}))$. Hence it follows from the assumptions on ℓ and Φ, having quadratic growth, that $\inf\{J(B), B \in \mathcal{B}_{ad}\} > -\infty$. Thus there exists an $m_0 \in R$ such that $\lim_{n\to\infty} J(B_n) = m_0 > -\infty$. This completes the proof. □

Acknowledgement

Thanks to Professors Matania Ben-Artzi, Simeon Reich and Alexander Zaslavski for their kind invitation to make this contribution to the special issue of PAFA dedicated to Professor Aizik Volpert.

References

[1] N. U. Ahmed, *Semigroup Theory With Applications to systems and Control*, Pitman research Notes in Mathematics Series, vol. 246, Longman scientific and technical, U.K (co-published in U.S by John Wiley & Sons. Inc, New York), 1991.

[2] N. U. Ahmed, *Linear and Nonlinear Filtering for Scientists and Engineers*, World Scientific, Singapore, New Jersey, London, Hong Kong, 1998.

[3] N. U. Ahmed, *Optimal control of infinite dimensional stochastic systems via generalized solutions of HJB equations*, Discussiones Mathematicae Differential Inclusions, Control and Optimization **21** (2001), 97–126.

[4] N. U.Ahmed, *Measure valued solutions for stochastic evolution equations on Hilbert spaces and their feedback control*, Discussiones Mathematicae Differential Inclusions, Control and Optimization **25** (2005), 129-157.

[5] N. U. Ahmed, *Weak Compactness in the space of operator valued measures $M_{ba}(\Sigma, \mathcal{L}(X, Y))$ and its applications*, Discussiones Mathematicae Differential Inclusions, Control and Optimization **31** (2011), 231–247.

[6] N. U. Ahmed, *Stochastic neutral evolution equations on Hilbert spaces and their partially observed relaxed control*, Journal of Abstract Differential Equations and Applications **5** (2014), 1–20.

[7] N. U. Ahmed, *Infinite dimensional uncertain dynamic systems on Banach spaces and their optimal output feedback control*, Differential Inclusions, Control and Optimization **35** (2015), 65–87.

[8] N. U. Ahmed, *Necessary conditions of optimality of output feedback control law for infinite dimensional uncertain dynamic systems*, Pure and Applied Functional Analysis **1** (2016), 159–184.

[9] B. Bongiorno, L.D. Piazza and K. Musial, *A characterization of the weak Radon-Nikodym property by finitely additive interval functions*, Bull. Aust. Math. Soc. **80**, (20019), 476–485.

[10] Z. Brzezniak and H. Long, *A Note on γ-Radonifying and Summing Operators*, Stochastic analysis, Banach center Publications, vol. 105, Institute of Mathematics, Polish academy of Sciences.

[11] Da Prato and J. Zabczyk, *A note on stochastic convolution, stochastic analysis and applications* **10** (1992), 143–153.
[12] J. Diestel and J. J. Uhl Jr., *Vector Measures, in: Mathematical Surveys*, vol. 15, American Mathematical Society, Providence, Rhode Island, 1977.
[13] N. Dunford and J. T. Schwartz, *Linear Operators Part 1, General Theory*, Inter science Publishers, Inc., New York, 1958,
[14] N. Ghoussoub and E. Saab, *On the weak Radon-Nikodym property*, Proc. of the American Mathematical Society **81** (1981), 81–84.
[15] S. Graf, *Selected results on measurable selections*, in: Proceedings of the 10th Winter School on Abstract Analysis, Circolo Matematico di Palermo, Palermo, Zdenêk Frolik (ed), Rendiconti del Circolo Matematico di Palermo, Serie II, Supplemento No. 2, 1982, pp. 87–122.
[16] S. Hu and N. S. Papageorgiou, *Handbook of Multi Valued analysis, Vol.1, Theory*, Kluwer Academic Publishers, Dordrecht, Boston, London, 1997.
[17] Y. Hu and S. Peng, *Adapted solutions of a backward semilinear stochastic evolution equation*, Stochastic Analysis and Applications **9** (1991), 445–459.
[18] L. Janicka, *Some Measure-theoretical characterization of Banach spaces not containing ℓ_1.*, Bull. Acad. Polon. Sci. Ser. Sci. Math. **27** (1979), 561–565.
[19] J. M. A. M. Van Neerven, M. C. Verrar and L. Weis, *Stochastic evolution equations in UMD Banach spaces*, Journal of Functional Analysis **255** (2008), 940–993.

Manuscript received August 21 2018
revised January 11 2019

N. U. Ahmed
University of Ottawa, Ontario, Canada
E-mail address: nahmed@uottawa.ca